1

2

Retrievals and Their Uncertainties:

What Every Atmospheric Scientist and Meteorologist Should Know

D.D. Turner[1], G.G. Mace[2], U. Löhnert[3], K. Ebell[3], and J.M. Comstock[4]

[1] National Severe Storms Laboratory / NOAA
[2] University of Utah
[3] University of Cologne
[4] Pacific Northwest National Laboratory

Corresponding Author:
Dr. David Turner
National Severe Storms Laboratory / NOAA
120 David L. Boren Blvd
Norman, OK 73072
Voice: +1-405-325-6804
Email: dave.turner@noaa.gov

Capsule Statement

Characterizing the uncertainty in a retrieved variable requires understanding the
uncertainties from three different sources.

1

Abstract

3

Remote sensors are heavily used to provide observations for both the operational and research communities. These sensors typically do not make direct observations of the desired geophysical variables, but instead retrieval algorithms are used to derive the desired variables from the observations. It is critically important that our community understands the underlying assumptions made by many retrieval algorithms, including that the retrieval problem is often ill-posed and that there are various sources of uncertainty that need to be treated properly. In short, the retrieval challenge is to invert a set of noisy observations to obtain estimates of geophysical quantities using imperfect forward models and imperfect prior knowledge, and the problem is often complicated by the existence of nonunique solutions.

There are three sources of uncertainties that contribute to the uncertainties of all retrieved data products: uncertainties in the observations, uncertainties in the dataset used to develop the retrieval algorithm or constrain it, and uncertainties in the model used to simulate the instrument observations (called a forward model). These three sources of uncertainties are discussed and shown how they can contribute to uncertainties in the retrieval.

21

1. Introduction

Science relies on observations to develop theories about nature, and ultimately to determine if those theories accurately approximate how nature works. In atmospheric sciences, these observations come from both our natural senses (e.g., our eyes) and from instruments that we have developed. The sustained development of advanced instrumentation continues to open new horizons in our understanding about how nature, including the multitude of processes in our atmosphere, really operates. In fact, instrument development and scientific advancement typically progress together. As a matter of fact, there is a strong association between the dearth of observations regarding the cycling of water through the atmosphere via clouds and precipitation and the difficulty the community now faces in accurately predicting climate change with cloud processes being identified as a leading cause of uncertainty in models of the climate system (Dufresne and Bony, 2008).

A wide variety of in-situ and remote sensing techniques are used to characterize, understand, and quantify properties and processes that occur in the atmosphere. Improving our understanding of these processes, and how they interact with each other and the environment, is critically important to improving our ability to represent these processes in models (both numerical weather prediction and climate models). For example, model developers have called on the cloud observational community to close the gap so that improved parameterizations can be implemented in climate models (Morrison and Gettleman, 2008).

1

2 In response to this recognized need, our field has seen an explosion in the number and

3 diversity of remote sensing instrumentation. We are using advanced active remote sensors

4 such as single- and multi-wavelength lidars, radars of various wavelengths and capabilities,

5 sodars, scintillometers, and Global positioning systems. We are using passive remote

6 sensors like infrared spectrometers, microwave radiometers, imaging radiometers that

7 operate at wavelengths from the visible to the infrared, and beyond. All of these

8 instruments are taking advantage of various physical laws, many embodied in the

9 principles of radiative transfer, to gain new insights into the processes in the atmosphere

10 that tend to move water through the system.

11

12 While we have been creative in using remote sensors to observe the atmosphere and

13 planet, there is a common thread that holds in most of these observations: we are not

14 actually observing what we want to know. These instruments are measuring a change in

15 voltage, the number of photons passing into a detector over a certain time period, a

16 Doppler shift in radar or lidar frequency, and the intensity of the backscattered energy.

17 What we often really desire to know is the geophysical variable: the ice water content at a

18 certain altitude or within a certain volume, the temperature profile, the aerosol and cloud

19 droplet number concentration, etc. Thus, we are left with the problem of extracting the

20 information we want from observations that are hopefully sensitive to the geophysical

21 quantity of interest. This inverse process is called a retrieval.

22

1   Very often, if we have a measurement that has some sensitivity to the geophysical variable

2   we desire, we can compute the signal that we would observe with our remote sensor using

3   a forward model (Fig 1) that begins with an atmospheric state that includes the geophysical

4   quantity of interest and reproduces a measurement after various assumptions and

5   approximations are made. These forward models ideally are based upon first-principles, so

6   that we have a fair degree of confidence in their fidelity. However, they are often nonlinear,

7   which makes them difficult if not impossible to invert analytically.   Thus, the retrieval

8   problem is essentially the development of an algorithm that is used to invert the forward

9   model (F) so that we can derive our geophysical variable that we desire ($\mathbf{X}$) from the

10   observation that we have made from our remote sensor ($\mathbf{Y}$). (Table 1 provides a brief

11   summary of the meaning of all symbols used in this paper.)

12

13   Graeme Stephens (1994), in his seminal book on remote sensing, provides a classic

14   illustration of the difference between the forward model and retrieval. Suppose that what

15   you desire is a description of a dragon, and what you observe are the footprints that the

16   dragon makes in the sand on a beach. Now, if you already know $\mathbf{X}$ (i.e., the dragon's weight,

17   size, how it moves, etc.) you can pretty easily describe the tracks it might make in the sand;

18   in other words, you can develop the forward model and predict the type of observations

19   that could be made – namely $F(\mathbf{X})$. But if you observe only the tracks in the sand, it will be

20   much more difficult to describe the dragon in any detail. You will likely be able to tell from

21   the tracks that the animal wasn't a deer, for example, by the shapes of the tracks, and

22   perhaps by the depth of the indentations say something about the size and weight of the

dragon. But there will be aspects about the dragon that you will be unable to state with any

certainty: if it is male or female, its color, if it has wings, what it ate for lunch, etc.

The retrieval process is complicated by many factors. First, there is the uniqueness

problem. There is no guarantee that there is only a single **X** that maps to the observation **Y**;

it is quite possible that $F(\mathbf{X}) = F(\mathbf{X}') = \mathbf{Y}$ where $\mathbf{X} \neq \mathbf{X}'$. This implies that there is a

distribution of states around **X** that will map to **Y**, but also the potential of two very

different states mapping to Y. Because of this and because there are often more unknowns

than there are measurements, inversion problems are more often than not ill-posed.

Second, there is no such thing as a perfect instrument; the measurements made by all

instruments have some noise component to them implying that the measurement, **Y**, can

also be represented by a probability distribution. Third, although the forward models are

often based on first principles, there are still uncertainties within the models: either in the

physics themselves (single scattering properties, for instance) or in the ancillary input

datasets that are needed to drive the model (along with the geophysical variable we desire)

to simulate the observation. Thus, we have to invert a set of noisy observations using an

imperfect forward model where multiple discrete descriptions of the atmospheric state

could reproduce the measurements!

A wide range of techniques has been developed to perform retrievals. There are a number

of good resources available that discuss retrieval theory (e.g., Twomey 1977, Tarantola

2005, Rodgers 2000, Stephens and Kummerow 2007; Aster et al. 2013), and many scientific

disciplines use inversion theory as part of their normal activities. Since retrieved data is

1   utilized so heavily in our field for both operational meteorology and research (e.g., satellite

2   estimates of temperature profiles and cloud properties, weather radar estimates of

3   precipitation amount and intensity, etc.), it is critical that users of such data streams have

4   an understanding of the uncertainties in the data products. Furthermore, as we continually

5   develop more advanced instrumentation and algorithms, it is important that the

6   uncertainties in these retrievals are accurately quantified and made available with the

7   retrieved information. Thus, our objective here is to discuss uncertainties in retrievals in

8   general terms, and suggest where improvements in the quantification of these

9   uncertainties need to occur.

10

11   ## 2. Background on Retrievals

12

13   Before we start discussing the sources of uncertainty in retrievals, we need to first start

14   with some basic background on retrieval methodology. For ease of discussion, we classify

15   these algorithms into two distinct groups: "regression-based" and "variational-based".

16

17   The regression-based retrieval algorithms develop empirical relationships between the

18   geophysical variable **X** and the observation **Y** using techniques like linear or polynomial

19   regression algorithms, empirical orthogonal functions, neural networks, and theoretical

20   relationships between scattering, emission, and transmission with the geophysical

21   properties of interest (e.g. Nakajima and King, 1990). These are essentially statistical

22   inferences. Often, a prior dataset of the geophysical observations (i.e., a climatology of

1   those observations) is used together with a forward model to simulate the observations

2   that would be associated with each **X**. The accuracy of a regression-based retrieval will

3   depend strongly on the statistical properties of the prior dataset (i.e., the mean $\mathbf{X_a}$ and the

4   covariance matrix of the prior dataset $\mathbf{S_a}$) used in the construction of the retrieval relative

5   to the particular case being retrieved. If the current case is well represented by the prior

6   dataset (i.e., **X** is "close" to $\mathbf{X_a}$) then the retrieval will likely be reasonably accurate, but if

7   the current case is poorly represented in the prior dataset (i.e., **X** is "far" from $\mathbf{X_a}$) then the

8   accuracy of the retrieval will likely be low, especially if the forward model is nonlinear. To

9   use a more concrete example: suppose that only data from the tropics were used to develop

10   a regression-based retrieval for a satellite sensor. The retrieved variable will likely be

11   much more accurate when applied to observations taken in the tropics and perhaps not

12   very accurate when applied to data collected in polar regions. Often regional data sets or

13   specific cloud resolving model output is used for this purpose. Examples of these

14   regression-based retrievals include algorithms that relate microwave brightness

15   temperature (e.g., Conner and Petty 1998) or infrared emission (e.g., Adler et al. 2003) to

16   rain rate, and algorithms that relate ice water content to millimeter radar reflectivity (e.g.,

17   Protat et al. 2007).

18

19   Variational-based retrieval algorithms exploit the idea, introduced earlier, that

20   observations, forward models, and the assumptions necessary to drive them have inherent

21   uncertainty and can be represented by probability distributions. These algorithms

22   typically use some application of Bayes Theorem (Eq 1), which states that the probability of

23   retrieving the geophysical variable **X** given the observation **Y** is equal to the probability of

1 making the observation **Y** given the current state of the geophysical variable **X** times the

2 probability that **X** occurs based upon some prior knowledge (i.e., a climatology, a model

3 forecast, or a measurement of **X** at some distinct distance away in time or space). Note that

4 the denominator in Eq 1 is a normalization factor.

5

6
$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)}$$
Eq 1

7

8 Thus, the retrieved value of **X** is also a statistical solution, but one that combines the prior

9 knowledge of **X** (i.e., the climatology) with the forward model $P(\mathbf{X|Y})$ into a maximum

10 likelihood solution.

11

12 Like regression-based retrievals, there are a wide variety of variational-based retrieval

13 algorithms. Many of these techniques are described by Rodgers (2000). Variational

14 methods often are trying to find an optimal solution that minimizes a cost function J (Eq 2)

15 to get a solution that agrees with (a) the measurements[†] within their uncertainties (**S$_e$**) and

16 with (b) the prior knowledge (**X$_a$**) within its uncertainties (**S$_a$**).

17

18
$$J = (Y - F(X))^T S_e^{-1}(Y - F(X)) + (X - X_a)^T S_a^{-1}(X - X_a)$$
Eq 2

[†] As will be shown later, the "measurements" include both the observations made by the instrument that the forward model is simulating and the uncertainties in the forward model itself.

1

2   Minimizing this cost function is often done in an iterative manner, wherein a guess of **X** is

3   compared to **Y** via the forward model and with the climatological mean $\mathbf{X_a}$; **X** is then

4   updated according to some mathematical approach. Typically, the update to **X** requires

5   knowing or computing the Jacobian of *F* with respect to **X** (i.e., the sensitivity of the forward

6   model to perturbations of the desired geophysical variable). The retrieval algorithm

7   continues to update **X** until the algorithm converges on a point that minimizes *J* (or the

8   algorithm has determined that it has diverged and no solution can be found).

9

10   An increasingly common approach being used in atmospheric science is optimal

11   estimation, which assumes that the uncertainties in the prior information, observation, and

12   model uncertainties are Gaussian and that the forward model is approximately linear

13   (Rogers 2000) and can be represented by the first derivative of **Y** with respect to **X** (i.e., by

14   the Jacobian of *F*). If the forward model is nearly linear (i.e. the higher order derivatives

15   can be neglected), then the forward model can be linearized around some prior state (e.g.,

16   $\mathbf{X_a}$), and this is adequate to find a solution. Moderately non-linear forward models (and in

17   atmospheric science many forward models are nonlinear due to the propensity of

18   exponentials and power laws used within them) can be used in this framework using

19   Gauss-Newton methods to minimize *J*, but techniques such as regularization and weighting

20   approaches (e.g., Carissimo et al. 2005, Turner and Löhnert 2013) can be used to help

21   stabilize the iterative search and prevent the algorithm from diverging.

22

1 An advantage of variational methods is that the uncertainties of the retrieved solution are

2 typically derived simultaneously since the probability distribution variance of the solution

3 can be taken to represent the uncertainty, whereas for most regression-based methods

4 additional equations and approaches need to be derived to estimate the uncertainties (e.g.,

5 Cadeddu et al. 2009). A disadvantage of the variational methods is that since they are

6 usually iterative algorithms, they are more expensive computationally than regression-

7 based methods. This is especially true if the Jacobian of F has to be determined numerically

8 for each iteration.

9

3. Sources of Retrieval Uncertainty

11

12 Every retrieval algorithm, whether regression based or variational, suffers from the same

13 fundamental sources of uncertainty such as observational error, forward model error, or

14 the ill posed nature of the retrieval process. The first, and perhaps most obvious, source of

15 uncertainty is the uncertainty in the observations themselves. Since the observation is

16 typically multi-dimensional (i.e., there are observations from multiple channels from a

17 single instrument, or multiple instruments used in the retrieval), there will be random

18 uncorrelated error for each particular element of **Y**, as well as correlated error between the

19 different elements of **Y**. This can be succinctly specified in a covariance matrix $\mathbf{S_y}$.

20

21 For many instruments, it is easier to estimate the uncorrelated error in each particular

22 element of the observation. For example, the basic measurement of lidars is typically the

1 number of photons per range bin, and thus the uncertainty in each bin is usually a Poisson

2 distributed random error. However, determining the correlated error in a measurement is

3 typically much harder but can be done; for example, Tobin et al. (2007) have used a

4 principle component analysis technique to estimate the correlated error in the

5 Atmospheric Infrared Sounder (AIRS) sensor that is flying on the Aqua satellite.

6

7 The second source of uncertainty is associated with the forward model. While forward

8 models are usually built on first principle understanding, many (necessary) estimations

9 and approximations still exist and are used either due to limitations in our understanding

10 of the physics or for computational efficiency reasons. The use of a 1-D radiative transfer

11 model instead of a 3-D radiative transfer model is one example of an approximation that is

12 frequently used, even though the 3-D effects are important in most cloud scenes (e.g.,

13 Pincus et al. 2012, Liang and Di Girolamo 2013, Zhang et al. 2013). Furthermore, the inputs

14 needed by F often consist of more than just the elements of **X**. These so-called model

15 parameters need to be known and their uncertainty will also contribute to the uncertainty

16 of F. These model parameters can be represented by a vector **B** and their uncertainty by

17 the covariance matrix $\mathbf{S_b}$.

18

19 The forward model uncertainties are typically ignored in most current retrieval algorithms;

20 this is equivalent to assuming that the forward model is perfect. However, this assumption

21 can result in a drastic underestimation of the error in the retrieved **X**, and for some cases,

22 ignoring the uncertainties associated with **B** is the Achilles' heel of the retrieval itself.

1   Posselt and Mace (2013) provide a good example of this phenomenon; another example is

2   shown in Fig 4 of Turner et al. (2007b).

3

4   Perhaps one reason why the uncertainties in the forward model are ignored is that many

5   algorithm developers associate $\mathbf{S_e}$ = $\mathbf{S_y}$ in Eq 2. However, we can easily translate the

6   uncertainties in $\mathbf{B}$ by the forward model into the observation space, and thus combine the

7   uncertainties of the observations and forward model together. This is done by computing

8   the Jacobian of F with respect to $\mathbf{B}$, denoted as $\mathbf{K_b}$, and then combining $\mathbf{S_y}$ and $\mathbf{S_b}$ using this

9   Jacobian to get the "total measurement uncertainty" $\mathbf{S_e}$ as

10

11                          $$S_e = S_y + K_b S_b K_b^T$$ Eq 3

12

13   A third source of uncertainty is associated with the prior dataset that is used to either build

14   the regression-based retrieval or serve as a constraint in the variational-based retrieval.

15   Within the climatology of the desired state vector $\mathbf{X}$, there will be a natural distribution of

16   values for each element in $\mathbf{X}$ as well as correlations between the different elements in $\mathbf{X}$.

17   Again, if we assume Gaussian statistics we can represent this uncertainty in the prior with a

18   covariance matrix $\mathbf{S_a}$ around the mean climatological state $\mathbf{X_a}$.

19

20   The prior dataset serves as a tremendous constraint for the entire retrieval; in fact, the

21   uncertainty in the retrieval must be smaller than the uncertainty in the prior otherwise the

22   retrieval does not add any information from the observations. However, if we do not have

23   a good measure of the variability and covariance of the true atmospheric state, then the

1   retrieval will not be able to use this information and could easily give a biased result. In

2   some cases, the prior does not serve as a serious constraint because the information

3   content in the observations is so high, such as retrieving liquid water path from a

4   microwave radiometer (e.g., Turner et al. 2007a).   However, in some cases the prior

5   covariance is critically important, such as retrieving liquid water content profiles from a

6   combination of microwave radiometer and cloud radar observations (e.g., Ebell et al.

7   2010). And unfortunately, there are some geophysical variables where the $\mathbf{S_a}$ matrix is not

8   well characterized by observations (e.g., such as the vertical level-to-level covariance of

9   liquid water content), and thus model simulations are used to provide this information in

10   order to serve as a constraint in the retrieval (Löhnert et al. 2001).

11

12   Thus, there are three covariance matrices that must be specified in order to characterize

13   the uncertainty in the retrieval: $\mathbf{S_y}$, $\mathbf{S_a}$, and $\mathbf{S_b}$. These covariance matrices are combined

14   together to get the uncertainty of the optimal solution $\mathbf{S_{op}}$

15

16
$$S_{op} = \left(S_a^{-1} + K_{op}^T S_e^{-1} K_{op} \right)^{-1} \qquad\qquad \text{Eq 4}$$

17

18   where $\mathbf{K_{op}}$ is the Jacobian of F with respect to the optimal solution $\mathbf{X_{op}}$.   Thus, the

19   uncertainty of the solution also includes a measure of the amount of sensitivity of the

20   forward model to changes in the desired geophysical variable. This is an important aspect

21   that needs to be captured because many instruments can experience saturation effects and

22   lose sensitivity for some portion of the range of $\mathbf{X}$.   For example, a cloud will become

23   opaque at a particular wavelength if the liquid water path (LWP) becomes large enough,

1   and thus the radiometer measuring at that wavelength no longer has sensitivity to changes

2   in LWP for LWP values above that threshold.   This is a particularly important but often

3   ignored issue for infrared and visible spectrum retrievals done either from the ground or

4   from space.

5

6   This discussion has centered upon random uncorrelated and correlated errors, which can

7   be represented in $\mathbf{S_y}$, $\mathbf{S_a}$, and $\mathbf{S_b}$. However, systematic errors can occur in both F (via $\mathbf{B}$ and

8   with assumptions such at 1-D radiative transfer) and $\mathbf{Y}$. The equations listed above will not

9   propagate systematic uncertainties into the solution.   It is critically important that

10   scientists developing instruments and forward models perform careful evaluations to

11   identify and remove systematic errors. Retrievals from a poorly calibrated instrument will

12   nearly always be suboptimal. One way to evaluate both F and $\mathbf{Y}$ for systematic errors

13   simultaneously is to perform closure studies.   A closure study of this type has three

14   components: the observations that are being evaluated, and forward model and its internal

15   parameters (i.e., the subset of $\mathbf{B}$ that are internal to the model), and the observations that

16   are needed to drive the forward model (i.e., the remainder of $\mathbf{B}$ and the geophysical

17   variables that are typically retrieved). Delamere et al. (2010) provide an example of an

18   infrared spectral radiance closure study. In some cases, the separation of $\mathbf{B}$ in this manner

19   results in a separation of the "known unknowns" (i.e., the parameters that are known to

20   have an impact on F but perhaps are uncertain) and the "unknown unknowns" (i.e., aspects

21   of F that are uncertain but that the user was unaware of). A well-structured closure study

22   should be able to identify if there are any "unknown unknowns" and also if there are

23   systematic errors in the "known unknowns", F, and $\mathbf{Y}$.

1

4. Simple Examples

3

4   We will use several examples to demonstrate how the uncertainties in $\mathbf{S_y}$, $\mathbf{S_a}$, and $\mathbf{S_b}$ impact

5   the uncertainties in $\mathbf{S_{op}}$. While we could have chosen from a wide range of retrievals used

6   in atmospheric science, we have chosen a simple example that is straightforward to

7   illustrate and hopefully easy to understand. In this problem, we fit a Gaussian-like curve,

8   defined by Eq 5, to an observed distribution of values over a fixed resolution axis z. Like

9   many forward models used in atmospheric science, this model is moderately nonlinear.

10

11 $$F_z = N \exp\left(-\left(\frac{z-R}{W}\right)^2\right) \qquad\qquad \text{Eq 5}$$

12

13   In our examples (Figs 2 and 3, and online supplementary figures S1-S4), we create sensor

14   measurements (i.e., observations) of a true distribution described by Eq 5 and then, for

15   various assumptions of the values in $\mathbf{S_y}$, $\mathbf{S_a}$, and $\mathbf{S_b}$, retrieve the parameters of the model N,

16   R, and W (i.e., our desired geophysical variables) that describe the true distribution as well

17   as the uncertainties in these parameters. To help visualize the solution, we will use F (i.e.,

18   Eq 5) to translate the retrieved solution back into the observational space. This example,

19   while simple, has practical applications also: for example, Miles et al. (2000) essentially

20   retrieved N, R, and W from a large set of in-situ cloud droplet size distribution observations

21   in stratus clouds.

1

2   The true distribution is given by N=300, R=60, and W=10 (black curves in panel A in Figs 2,

3   3, and S1-S4). The observation (green curves) was created from this true distribution by

4   applying a uniform random error (i.e., same uncertainty for each bin element z) and

5   assuming that the errors in the observations are uncorrelated; the same observation will

6   be used in all of these examples. These examples are designed to illustrate how changes in

7   **S_y, S_a,** and **S_b,** where the latter implies uncertainties in *F*, impact the solution **X_{op}** and its

8   uncertainty **S_{op}.**

9   *a) Perfect model*

10   The first example illustrates a "perfect model" case where the state vector **X**=[N, R, W] and

11   there are no unknown parameters in *F*. The assumed prior distribution **X_a**=[200, 50, 15],

12   with 1-σ uncertainties of 200, 20, and 5, respectively. We assumed that there was no

13   correlation between N, R, and W in the prior, resulting in

14

15
$$S_a = \begin{bmatrix} \sigma_N^2 & 0 & 0 \\ 0 & \sigma_R^2 & 0 \\ 0 & 0 & \sigma_W^2 \end{bmatrix} = \begin{bmatrix} 200^2 & 0 & 0 \\ 0 & 20^2 & 0 \\ 0 & 0 & 5^2 \end{bmatrix}.$$

16

17   Fifty observations where made of this distribution (green curve), with an observation at

18   every even value of z. The assumed 1-σ uncertainty in the observations is 100; thus **S_y** is a

19   diagonal matrix with values of 100^2 along the diagonal. Since *F* is a perfect model, **S_b**=0

20   resulting in **S_e**=**S_y**. The retrieval used an optimal estimation approach (Rodgers 2000) to

1   minimize the cost function J (Eq 2) and derive $\mathbf{X_{op}}$, with the uncertainty of the solution

2   computed using Eq 4.

3

4   The retrieved solution in observation space (i.e., as $F(\mathbf{X_{op}})$) is shown in Fig 2A, with the

5   retrieved values and the 1-σ uncertainties listed in the inset black box (along with the truth

6   and assumed prior values and uncertainty). In this case, the retrieved values are N=266.9 ±

7   43.4, R=60.8 ± 1.6, and W=11.4 ± 2.0 where the 1-σ values were computed from the square

8   root of the diagonal elements of $\mathbf{S_{op}}$. This successful retrieval shows the retrieved values

9   agreeing with the truth within the uncertainties of the retrieval.

10

11   However, there is also some correlation in the uncertainty of the different elements in the

12   retrieved $\mathbf{X_{op}}$: the Pearson correlation coefficient r between N and W is equal to -0.53 and is

13   due to off-diagonal elements in the Jacobian $\mathbf{K_{op}}$ (i.e., due to the forward model

14   correlations). The correlated error between the other pairs of parameters (N and R, and R

15   and W) is zero; again, this lack of correlation is due to the forward model. Since $\mathbf{X_{op}}$

16   specifies the most probable solution between the observations and the prior, it is useful to

17   visualize the range of solutions that lie within its uncertainty $\mathbf{S_{op}}$. We did this by sampling

18    $\mathbf{S_{op}}$ (as well as $\mathbf{S_a}$ and $\mathbf{S_e}$) using a Monte Carlo technique (Tarantola 2005), and then

19   translating these results back into the observational space with F (Fig 2B). The "flattening"

20   (or the plateau) of the blue trace in Fig 2B is due to the correlation between N and W in $\mathbf{S_{op}}$.

21   Also, note that the uncertainty in the solution is much smaller than the uncertainty in the

22   prior; in the case where the observations add no information on the state vector then the

23   uncertainty in the solution would be the same as the prior's uncertainty. The online

1 supplemental material illustrates how the solution and its uncertainties change as the

2 magnitude of uncertainty in the observations (i.e., $\mathbf{S_y}$) changes (example S1).

3 ### b) Imperfect model

4 The previous example assumed that the model is perfect and that there are no uncertain

5 parameters within the model. However, this is not typical as forward models often are

6 quite complex. Thus, we will assume that one of the three parameters in our forward

7 model (i.e., one of N, R, and W) is now a model parameter with some uncertainty.

8

9 We will assume that W is the model parameter and fix its value to W=15±5. Thus, the true

10 W is within the uncertainty of the model parameter, but they are not the same (i.e., the

11 model has some bias). This implies that $\mathbf{S_b}$ = 5^2 (i.e., a scalar in this case), and thus $\mathbf{S_e}$ will

12 have contributions from both $\mathbf{S_y}$ and $\mathbf{S_b}$ (Eq 3). The retrieval will then solve for $\mathbf{X}$=[N, R]. In

13 this example, the assumed $\mathbf{S_y}$ is 4x smaller than that shown in Fig 2, but is the same as the

14 S1 case in the online supplemental section.

15

16 The retrieved solution and its uncertainty are shown in Fig 3. There are several differences

17 between the solution here and the perfect model solution (Fig 2). First, it is clear that the

18 derived distribution (i.e., $F(\mathbf{X_{op}})$) for the imperfect model does not agree as well with the

19 truth as when the perfect model was used in the retrieval. This will be true for all cases

20 when a biased forward model is used. However, the retrieved values of N and R do agree

21 within 2-σ of the true values.

22

However, there is also a striking difference in the combined observational and forward model uncertainty (green curve, derived from S_e) between Fig 2B and Fig 3B. The "enhancement" to this total measurement uncertainty (i.e., its double-peaked feature, relative to its shape in Fig 1) is from the uncertainty in the forward model parameter W. Note that as the state vector evolves during the iterative retrieval, the observational uncertainty also evolves; this is because the Jacobian with respect to B (i.e., K_b in Eq 3) changes as X changes. This results in a more broadly distributed uncertainty in S_{op} (blue curve in Fig 4B) that has a less pronounced plateau than that exhibited in Fig 2B.

Additional examples using an imperfect model that show the impacts of a change in the prior uncertainty (S2), a different imperfect model (S3), and using a significantly smaller number of observations in the retrieval (S4) are included in the online supplement.

5. Example Retrieving Liquid Water Content

The examples above and in the online supplement illustrate how changes in the uncertainties in the observations, prior dataset, and forward model impact the uncertainties in the retrieval; however, the forward model/retrieval problem was highly simplified to illustrate these points. Here, we use a real-world retrieval problem to illustrate these same points.

Millimeter-wave cloud radars have been deployed at many locations worldwide to provide observations of macro- and microphysical cloud properties that can be used to develop

1   cloud climatologies, evaluate and improve processes within cloud-resolving model

2   simulations, provide inputs needed for radiative transfer calculations, and more. One of

3   the primary observations from these radars is radar reflectivity (Z), and many retrieval

4   algorithms have been developed to retrieve profiles of liquid water content (LWC) from the

5   observed profiles of Z. We will use the simplified approach of Brandau et al. (2010) in a

6   variational framework to evaluate how uncertainties in $\mathbf{S_a}$, $\mathbf{S_y}$, and $\mathbf{S_b}$ impact the

7   uncertainty in the retrieved LWC (i.e., $\mathbf{S_{op}}$).

8

9   The Brandau algorithm assumes that the cloud droplet size distribution is mono-modal

10   with a generalized gamma size distribution, and that the cloud is homogeneously mixed

11   and thus the number concentration of cloud droplets (N_d) is constant with height. Their

12   forward model computes LWC (at some height) from Z using

13

14   $$LWC = \frac{\pi \rho_w}{6} \left(\frac{v(v+1)(v+2)}{(v+3)(v+4)(v+5)} \right)^{\frac{1}{2}} N_d^{\frac{1}{2}} Z^{\frac{1}{2}}$$ Eq 6

15

16   where ρ_w is the density of liquid water and v describes the slope of the gamma size

17   distribution. In this example, we treat N_d and v as model parameters and retrieve LWC

18   from Z.

19

20   Since the errors depend on the particular state of the atmosphere, we assume that the true

21   atmospheric state has LWC=0.1 g m^{-3}, N_d=500 cm^{-3}, and v = 8.7. [The value for v was

22   determined by Brenguier et al. (2011) from in situ measurements in stratus clouds, and has

1   an uncertainty of approximately 3.]   Using these values of *LWC*, N_d, and *v* yield a radar

2   reflectivity value of -37 dBZ, which is very typical of non-drizzling stratiform warm clouds.

3

4   Figure 4 shows the impacts of 1-σ uncertainties in N_d, *Z*, and **S_a** on the uncertainty in *LWC*.

5   For this example, we've assumed **X_a**=0.1 g m^{-3} and that the 1-σ uncertainty in this prior is

6   both 100% and 30% (open squares and filled circles, respectively). The uncertainty in the

7   retrieved *LWC* is computed over a range of uncertainties in *Z* (from perfect observations

8   with no uncertainty in red to relatively poor observations with 1-σ uncertainty of 3 dBZ in

9   brown). In these retrievals, the uncertainty in the model parameter N_d is changed from

10   perfect knowledge (0 cm^{-3}) to 60% error (300 cm^{-3}).   Figure 4 demonstrates that the

11   uncertainty in the forward model parameter N_d dominates when the uncertainty in *Z* is

12   small, but that as the uncertainty in the *Z* observations increases it begins to dominate the

13   uncertainty in the retrieved *LWC* relative to the uncertainty in N_d.   However, if the

14   uncertainty in the prior value of *LWC* is smaller (filled circles), then the uncertainty in N_d

15   becomes less important. Thus the importance of a source of uncertainty depends on the

16   other two sources of uncertainties in the retrieval.

17

18   **6. A Classification Conundrum**

19

20   There is a fourth source of uncertainty that affects many retrieval algorithms.   Many

21   retrieval algorithms are designed for particular conditions; for example, single-layer ice-

22   only clouds, non-drizzling clouds, dust-laden aerosol layers, etc.   This is often because

1   either (a) the forward model is accurate only in these particular conditions or (b) the

2   forward model is very sensitive to other geophysical variables but that there is either poor

3   constraint in the prior on these variables or they are highly correlated with other elements

4   in the state vector.

5

6   To overcome this limitation, the scene must be classified to determine if the conditions are

7   appropriate for the forward model assumptions used in the given retrieval algorithm.

8   These classification algorithms use many of the same datasets used in the retrieval

9   algorithms themselves, and indeed are just a form of a retrieval algorithm. For example,

10   there has been a wide range of algorithms developed (e.g., Illingworth et al. 2007, Pavolonis

11   et al. 2005, Shupe 2007). However, the output of these algorithms is typically discrete; for

12   example, in the case of the Cloudnet algorithm each radar / lidar range bin is classified as

13   an aerosol layer, a precipitation layer, liquid cloud layer, ice layer, etc. (Illingworth et al.

14   2007).

15

16   In some cases, classification algorithms may assign probabilities instead of discrete values.

17   For example, the cloud mask algorithm of Frey et al. (2008) uses four confidence levels to

18   indicate the likelihood of a MODIS pixel being non-clear (i.e., some obscuration in the field

19   of view by cloud). The phase algorithm of Riedi et al. (2010) provides a continuous index

20   to describe cloud phase likelihood using combined MODIS and POLDER observations.

21

22   Regardless, classification assignments need to be included as an importance source of

23   uncertainty in the subsequently retrieved geophysical variables. Therefore, methods need

1 to be developed to propagate classification uncertainties through the retrieval algorithm

2 itself. Propagating uncertainties in discrete variables that are used as input to another

3 retrieval algorithm requires some careful thought and evaluation. One possible way

4 forward is to incorporate the classification and cloud retrieval into a single algorithm;

5 another option is to use Monte Carlo techniques to propagate the uncertainties in the

6 discrete classification algorithms into the final geophysical variables. In either regard, the

7 uncertainty in the scene classification and the potential to apply a retrieval algorithm to a

8 scene for which it was not applicable is a huge source of uncertainty in the retrieval itself.

9 **7. Conclusions**

10 In summary, most of the observations we use to view the Earth-atmosphere system are

11 from remote sensors and these observations are used in a wide range of both operational

12 and research applications. These sensors include all satellite observations, weather radar

13 observations and many more. In all of these cases, geophysical variables are retrieved from

14 the observations using an algorithm that inverts, in some sense, the measurements that are

15 a function of the geophysical parameters to arrive at an approximation of those parameters

16 of interest. The uncertainties in these retrieved quantities arise from three sources: the

17 climatological (or prior) dataset used to either construct or constrain the retrieval, the

18 uncertainties in the forward model assumptions and input ancillary datasets used in the

19 forward model, and the uncertainties in the observations themselves. The covariance

20 matrices $\mathbf{S_a}$, $\mathbf{S_b}$, and $\mathbf{S_y}$ characterize these three sources of uncertainty, respectively. These

21 uncertainties, together with the sensitivity of the forward model to perturbations in the

22 model parameters $\mathbf{B}$ and the atmospheric state $\mathbf{X}$, can be propagated to provide

1   uncertainties in the retrieved atmospheric state $\mathbf{X_{op}}$ as the covariance matrix $\mathbf{S_{op}}$. We have

2   utilized the variational retrieval method to propagate these uncertainties in a series of

3   examples to show the impact of how changes in these uncertainties result in changes in the

4   retrieved state vector $\mathbf{X_{op}}$ and $\mathbf{S_{op}}$. These same conclusions can also be shown using

5   regression-based retrieval methods (e.g., Löhnert and Maier 2012).

6

7   We contend that more work is needed to accurately characterize these three input

8   covariance matrices. Too often, our community assumes that the measurement covariance

9   matrix $\mathbf{S_y}$ is diagonal and that there is no correlation between different measurements

10   within the observational vector $\mathbf{Y}$. We challenge instrument developers to devise methods

11   to test this assumption, and for scientists that develop retrievals to include the results from

12   these tests in their algorithms.

13

14   The prior dataset is absolutely critical for the development of regression-based retrievals

15   and to serve as constraints in variational-based retrievals. However, in many cases, the

16   climatology used to develop $\mathbf{X_a}$ and $\mathbf{S_a}$ is inadequate or must be estimated from other

17   sources such as model simulations. For example, ice crystal mass-dimensional and area-

18   dimensional relationships are a critical component for many ice cloud retrieval algorithms

19   (as well as for parameterizations in cloud resolving, weather, and climate models).

20   However, due to recently discovered shattering of ice crystals by the in-situ probes (e.g.,

21   Korolev and Isaac 2005, Lawson 2011), it has been argued that many of the older in-situ

22   datasets have biases that cannot be removed and thus an extensive new dataset needs to be

23   collected in order to define $\mathbf{S_a}$. Furthermore, as these ice particle properties are likely

1   correlated with height (temperature), these new datasets should be collected in a manner

2   to quantify these correlations so they can be captured within $\mathbf{S_a}$. There are many other

3   examples where improved prior datasets would greatly benefit the retrieval community.

4   We challenge the in situ observational community to analyze their data in such a way as to

5   define correlations between and among the various parameters they observe and derive.

6   Without these correlations, the retrieval community is left to either assume what these

7   correlations are or to ignore the correlations entirely.

8

9   Lastly, we believe that the retrieval community has neglected the uncertainties and

10   assumptions in the forward model parameters for too long. Forward models may be

11   fundamentally incorrect (e.g., applying 1-D radiative transfer approaches to scenes that are

12   inherently 3-D), or may have uncertainties with a parameter within the model. The

13   potential impact of uncertainty in the forward model parameters was illustrated in the

14   examples shown here. In some applications, the uncertainty in $\mathbf{B}$ dominates the

15   uncertainty in the observations, and potentially of the retrieval. This is especially true in

16   ice cloud and ice precipitation retrievals where particle mass and area are a function of ice

17   crystal habit. Realistic uncertainties in mass and area-dimensional relationships ($\mathbf{S_b}$)

18   dominate over the observational error ($\mathbf{S_y}$) in the total measurement uncertainty ($\mathbf{S_e}$), and

19   result in very large uncertainties in the retrieval (Posselt and Mace 2013). Characterizing

20   the uncertainties in F may require additional in-situ data (much in the same spirit as

21   improving the prior datasets) and supporting data so that closure studies can be

22   performed. Thus, a well-constructed set of field campaigns may be able to improve the

23   characterization of both $\mathbf{S_a}$ and $\mathbf{S_b}$ simultaneously.

Acknowledgments

References

Adler, R. F., and Coauthors, 2003: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.*, **4**, 1147–1167.

Aster, R.C., B. Borchers, and C.H. Thurber, 2013: *Parameter Estimation and Inverse Problems*. Elsevier, Oxford, 360 pp.

Brandau, C.L., H.W.J. Russchenberg, and W.H. Knap, 2010: Evaluation of ground-based remotely sensed liquid water cloud properties using shortwave radiation measurements. *Atmos. Res.*, **96**, 366-377, doi:10.1016/j.atmosres.2010.01.009.

1   Brenguier, J.-L., F. Fernet, and O. Geoffroy, 2011: Cloud optical thickness and liquid water

2   path – does the k coefficient vary with droplet concentration? *Atmos. Chem. Phys.*, **11**,

3   9771-9786, doi:10.5194/acp-11-9771.2011.

4   Cadeddu, M.P., D.D. Turner, and J.C. Liljegren, 2009: A neural network for real-time

5   retrievals of PWV and LWP from Arctic millimeter-wave ground-based observations.

6   *IEEE Trans. Geosci. Remote Sens.*, **47**, 1887-1900, doi:10.1109/TGRS.2009.2013205.

7   Carissimo, A., I. De Feis, and C. Serio, 2005: The physical retrieval methodology for IASI:

8   The δ-IASI code. *Environ. Model. Software*, **20**, 1111-1126.

9   Conner, M. D., and G. W. Petty, 1998: Validation and intercomparison of SSM/I rain-rate

10   retrieval methods over the continental United States. *J. Appl. Meteor.,* **37**, 679–700.

11   Delamere, J.S., S.A. Clough, V. Payne, E.J. Mlawer, D.D. Turner, and R. Gamache, 2010: A far-

12   infrared radaitive closure study in the Arctic: Application to water vapor. *J. Geophys.*

13   *Res.,* **115**, D17106, doi:10.1029/2009JD012968.

14   Dufresne, J.-L. and S. Bony, 2008: An assessment of the primary sources of spread of global

15   warming estimates from coupled atmosphere–ocean models. *J. Climate*, **21**, 5135–5144.

16   Ebell, K., U. Löhnert, S. Crewell, and D.D. Turner, 2010: On characterizing the error in

17   remotely sensed liquid water content profile. *Atmos. Res.*, **98**, 57-68,

18   doi:10.1016/j.atmosres.2010.06.002.

19   Frey, R.A., S.A. Ackerman, Y. Liu, K.I. Strabala, H. Zhang, J.R. Key, and X. Wang, 2008: Cloud

20   detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5. *J.*

21   *Atmos. Oceanic Technol.*, **25**, 1057-1072, doi:10.1175/2008JTECHA1052.1.

1 Illingworth, A.J., and coauthors, 2007: CloudNet – A continuous evaluation of cloud profiles

2 in seven operational models using ground-based observations. *Bull. Amer. Meteor. Soc.*,

3 **88**, 883-898.

4 Korolev, A.V., and G.A. Isaac, 2005: Shattering during sampling by OAPs and HVPS. Part I:

5 Snow particles. *J. Atmos. Oceanic Technol.*, **22**, 528-542.

6 Lawson, R.P., 2011: Effects of ice particles shattering on the 2D-S probe. *Atmos. Meas.*

7 *Technol.*, **4**, 1361-1381, doi:10.5194/amt-4-1361-2011.

8 Liang, L., and L. Di Girolamo, 2013: A global analysis on the view-angle dependence of

9 plane-parallel oceanic liquid water cloud optical thickness using data synergy from

10 MISR and MODIS. *J. Geophys. Res.*, in press, doi:10.1029/2012JD018201.

11 Löhnert, U., S. Crewell, A. Macke, and C. Simmer, 2001: Profiling cloud liquid water by

12 combining active and passive microwave measurements with cloud model statistics. *J.*

13 *Atmos. Oceanic Technol.*, **18**, 1354-1366.

14 Löhnert, U., and O. Maier, 2012: Operational profiling of temperature using ground-based

15 microwave radiometery at Payerne: Prospects and challenges. *Atmos. Meas. Tech.*, **5**,

16 1121-1134, doi:10.5194/amt-5-1121-2012.

17 Miles, N.L, J. Verlinde, and E.E. Clothiaux, 2000: Cloud droplet size distributions in low-level

18 stratiform clouds. *J. Atmos. Sci.*, **57**, 295-311.

19 Morrison, H. and A. Gettelman, 2008: A New Two-Moment Bulk Stratiform Cloud

20 Microphysics Scheme in the Community Atmosphere Model, Version 3 (CAM3). Part I:

21 Description and Numerical Tests. *J. Climate*, **21**, 3642–3659

1   Nakajima, T., and M. D. King, 1990: Determination of the optical thickness and effective

2   radius of clouds from reflected solar radiation measurements. Part I: Theory. *J. Atmos.*

3   *Sci.*, **47**, 1878–1893.

4   Pavolonis, M.J., A.K. Heidinger, and T. Uttal, 2005: Daytime global cloud typing from AVHRR

5   and VIIRS: Algorithm description, validation, and comparisons. *J. Appl. Meteor.*, **44**, 804-

6   826.

7   Pincus, R., S. Platnick, S.A. Ackerman, R.S. Hemler, and R.J.P. Hofmann, 2012: Reconciling

8   simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument

9   simulators. *J. Climate*, **25**, 4699-4720, doi:10.1175/JCLI-D-11-00267.1

10   Posselt, D. J., and G. G. Mace, 2013: MCMC-based assessment of the information content of a

11   combined radar and microwave cloud property retrieval. To be submitted to *J. Appl.*

12   *Meteor. Clim*.

13   Protat, A., J. Delanoe, D. Bouniol, A. J. Heymsfield, A. Bansemer, P. Brown, 2007: Evaluation

14   of ice water content retrievals from cloud radar reflectivity and temperature using a

15   large airborne in situ microphysical database. *J. Appl. Meteor. Clim.*, **46**, 557- 572.

16   Riedi, J., B. Marchant, S. Platnick, B.A. Baum, F. Thieuleux, C. Oudard, F. Parol, J.-M. Nicolas,

17   and P. Dubuisson, 2010: Cloud thermodynamic phase inferred from merged POLDER

18   and MODIS data. *Atmos. Chem. Phys.*, **10**, 11851-11865, doi:10.5194/acp-10-11851.

19   Rodgers, C.D., 2000: *Inverse methods for atmospheric sounding: Theory and practice*. Series

20   on Atmospheric, Oceanic and Planetary Physics, Vol 2, World Scientific, Singapore, 238

21   pp.

22   Shupe, M.D., 2007: A ground-based multisensory cloud phase classifier. *Geophys. Res. Lett.*,

23   **34**, 22, doi:10.1029/2007GL031008.

1 Stephens, G., 1994: *Remote Sensing of the Lower Atmosphere: An Introduction*. Oxford

2   University Press, New York, 523 pp.

3 Stephens, G. L., and C. D. Kummerow, 2007: The remote sensing of clouds and precipitation

4   from space: A Review. *J. Atmos. Sci.*, **64**, 3742-3763.

5 Tarantola, A., 2005: *Inverse Problem Theory and Methods for Model Parameter Estimation*.

6   Society of Industrial and Applied Mathematics, Philadelphia, PA, 342 pp.

7 Tobin, D.C., P. Antonelli, H.E. Revercomb, S. Dutcher, D.D. Turner, J.K. Taylor, R.O. Knuteson,

8   and K. Vinson, 2007: Hyperspectral data noise characterization using principle

9   component analysis: Application to the Atmospheric Infrared Sounder. *J. Atmos. Remote*

10   *Sens.*, **1**, 013515, doi:10.1117/1.2757707.

11 Turner, D.D., S.A. Clough, J.C. Liljegren, E.E. Clothiaux, K. Cady-Pereira, and K.L. Gaustad,

12   2007a: Retrieving liquid water path and precipitable water vapor from Atmospheric

13   Radiation Measurement (ARM) microwave radiometers. *IEEE Trans. Geosci. Remote*

14   *Sens.*, **45**, 3680-3690, doi:10.1109/TGRS.2007.903703.

15 Turner, D.D., and coauthors, 2007b: Thin liquid water clouds: Their importance and our

16   challenge. *Bull. Amer. Meteor. Soc.,* **88**, 177-190.

17 Turner, D.D., and U. Löhnert, 2013: Information content and uncertainties in

18   thermodynamic profiles retrieved from the ground-based Atmospheric Emitted

19   Radiance Interferometer (AERI). *J. Appl. Meteor. Clim.*, submitted.

20 Twomey, S., 1977: *Introduction to the Mathematics of Inversion in Remote Sensing and Direct*

21   *Measurements*. Elsevier, New York, 243 pp.

22 Zhang, Z., A.S. Ackerman, G. Feingold, S. Platnick, R. Pincus, and H. Xue, 2013: Effects of

23   cloud horizontal inhomogeneity and drizzle on remote sensing of cloud droplet

1   effective radius: Case studies based on large-eddy simulations. *J. Geophys. Res.,* in press,

2   doi:10.1029/2012JD017655.

3

1 **Online Supplemental Material**

2

3 We have included a few additional cases that use the Gaussian distribution forward model

4 (Eq 5) to illustrate how changes in the forward model and the assumed uncertainties can

5 impact the retrieval and its uncertainty. The IDL code that generated these examples, as

6 well as the examples in the paper, are available by request from the first author.

7   **i) Perfect model with Decreased Observational Uncertainty**

8 So what happens if the noise in the observations is decreased (i.e., S_y values are reduced) in

9 the perfect model case? To evaluate this, we decreased the assumed noise level in S_y by a

10 factor of 4 but all other parameters in the retrieval were the same as those shown in Fig 2.

11 The retrieved X_{op} (Fig S1) still agrees with the true parameters within the 1-σ uncertainties

12 of the retrieval; however, these uncertainties are now much smaller than the case shown in

13 Fig 2. This is shown both in the 1-σ values in the inset box as well in as the much smaller

14 values of the solution uncertainty (S_{op}), which was transferred back into observational

15 space as shown in Fig S1-B. Again, there is correlated error between the retrieved N and W

16 (r = –0.57), which results in the flat top of the posterior uncertainty trace (blue curve in Fig

17 S1-B).

18   **ii) Retrieving N and R with a modified prior and an imperfect model**

19 Suppose, however, that the uncertainty in the prior (i.e., S_a) is changed – what is the impact

20 on the retrieval? This change in the prior may result (for example) because the original

21 prior was determined from an annual climatology, and a decision was made to use data for

1   a specific month in the prior in order to reduce the uncertainty in the constraint. This case

2   is identical to the imperfect model case in the paper (i.e., Fig 3) except that the 1-σ value for

3   R in the prior is 1.0 instead of 20.0.

4

5   Comparing these results (Fig S2) with those from the original prior (Fig 3) shows a very

6   noticeable change in the distribution of the prior and its uncertainty. However, the impact

7   on the retrieved $\mathbf{X_{op}}$ is small, with $F(\mathbf{X_{op}})$ shifting leftwards towards smaller bin values (Fig

8   S2-A). The retrieved value of N is essentially unchanged, but the retrieved value of R is

9   shifted slightly towards the mean value of the prior (50) and is no longer within 2-σ of the

10   true value of R. This shifting is due to the algorithm optimizing the solution to minimize

11   the cost function J within the uncertainties of both the prior and the observations, and since

12   there is less uncertainty in the prior than the previous example it has more "pull". There

13   are also small changes in the uncertainty of the retrieval (Fig S2-B), with the uncertainty

14   also shifted slightly leftwards towards smaller bin values.

15

16   If an even smaller uncertainty in $\mathbf{S_a}$ was specified, the effect would be larger. If, in the

17   extreme, the 1-σ uncertainty in R used in $\mathbf{S_a}$ was set extremely small[†] (e.g, several orders of

18   magnitude smaller than the value), then the retrieval would not be able to modify this value

19   and it would be effectively fixed to the prior's value specified in $\mathbf{X_a}$. This, in essence,

---

[†] The 1-σ uncertainty for any element in $\mathbf{X_a}$ and $\mathbf{Y}$ (i.e., the diagonals of $\mathbf{S_a}$ and $\mathbf{S_y}$) must be

larger than zero, otherwise the associated matrices $\mathbf{S_a}$ and $\mathbf{S_e}$ would be ill-defined and the

inverse could not be computed.

1   reduces the dimensionality of the retrieval and can be quite useful if other observations

2   provide a tight constraint on one or more elements of the state vector. However, to

3   properly compute the error in the retrieval it is important that the uncertainty in the prior,

4   which is capturing the variability of this geophysical variable in nature, be properly

5   specified within the framework of the retrieval otherwise it is possible to over- or under-

6   constrain the retrieval resulting in a biased $\mathbf{X_{op}}$ and $\mathbf{S_{op}}$.

7   **iii) Retrieving N and W with the original prior and an imperfect model**

8   The imperfect model examples shown in section 4.b of the paper and case (ii) above

9   retrieved both N and R, which had uncorrelated uncertainties in the posterior (i.e.,

10   solution) in the perfect model cases (i.e., section 4.a and case (i) above). The uncertainties

11   in these variables were also uncorrelated in the imperfect model examples. However, the

12   uncertainties in N and W were correlated in the perfect model case, so this example will

13   use R as a model parameter and retrieve N and W.

14

15   Figure S3 shows the retrieved solution and the uncertainties when we assume R is the

16   model parameter with a value of 50±20 (as was the same as the uncertainty in the prior for

17   R in the previous examples). This example shows the poorest agreement between the

18   solution and the truth, with $F(\mathbf{X_{op}})$ not capturing the peak of the true distribution nor its

19   placement in the proper z bin, because the forward model was unable to move the peak of

20   the distribution with R held as a model parameter. Thus, the retrieved values of N and W

21   do not agree with the truth at all, yet the retrieval was considered converged within the

22   uncertainties of the observation and the prior. This is largely due to the much larger

1   combined observational and forward model error ($\mathbf{S_e}$), which is due to the large inflation

2   that resulted from the model parameter uncertainty $\mathbf{S_b}$. The posterior uncertainty also has

3   the largest spread in bin number (i.e., along the z-axis, with non-negligible values from bin

4   15 to bin 85) although the magnitude of the uncertainty isn't very large. Again, there is a

5   reasonably large Pearson correlation coefficient between the retrieved N and W, with $\mathbf{S_{op}}$

6   suggesting that r = −0.57.

7

8   This example also demonstrates conclusively the importance of evaluating, and potentially

9   improving, forward models relative to observations. In this case, there is a significant bias

10 between the true value of R and the value used in the forward model. Systematic errors

11 such as this are unable to be quantified by the covariance matrices: $\mathbf{S_a}$, $\mathbf{S_b}$, $\mathbf{S_y}$, and

12 consequently $\mathbf{S_{op}}$ only captures random and correlated errors.

13

14   **iv) Retrieving N and R using fewer observations using an imperfect model**

15 In all of the previous examples (Fig 2, 3 and S1-S3), the number of observations was 50,

16 which is much larger than the number of variables that were being retrieved (2 or 3, in

17 these examples). Thus, these examples are cases where the retrieval is over-constrained,

18 especially since the observational errors are uncorrelated and have good sensitivity to

19 changes in the forward model. However, it is much more common for the problem to be

20 under-constrained (often severely so) where the number of independent pieces of

21 information in the observations is less than the number of variables we are trying to

22 retrieve. For example, it is typical to use hundreds of spectral infrared observations to

1   retrieve temperature profiles at ~50 atmospheric levels, yet there may only be a few

2   independent pieces of information in the infrared spectra – certainly much fewer than the

3   number of levels that are being retrieved (e.g., Turner and Löhnert 2013).

4

5   To illustrate a more realistic retrieval, we reduce the number of observations taken of the

6   true size distribution from 50 to 4, where we only sample at bins 0, 33, 66, and 99 instead

7   of every even numbered bin along the *z*-axis. We then retrieve N and R using the imperfect

8   model where W is a model parameter.   This is essentially the same retrieval shown in

9   section 4.b in the paper, only where fewer observations were used. The results (Fig S4)

10 demonstrate much larger bias and uncertainties in the retrieved values of N and R when

11 fewer observations are used in the retrieval; for example, the bias error in the retrieved N

12 is nearly 30% with an uncertainty in the retrieved N of over 60%. Furthermore, there was

13 no correlated error between the retrieved N and R when 50 observations were used, but

14 when only 4 observations are used in the retrieval the correlated error between N and R is

15 −0.82. This significant change in the error in the retrieved solution is because there are

16 only 1.5 degrees of freedom of signal[†] in this example, which implies that there is less

17 information in the observations (1.5) than we are trying to retrieve (2; N and R).

18

19

[†] The degrees of freedom of signal (DFS), which is equivalent to the number of independent

pieces of information in the observations, is the trace of the averaging kernel **A**, where **A** is

derived from the **S_op, K_op,** and **S_e** matrices as $A = S_{op}^{-1} K_{op}^{T} S_{e}^{-1} K_{op}$ (Rodgers 2000).

1
2   **Table 1**: Summary of the symbols used

Notation	Meaning	Notes
Vectors		
Y	Observations (m elements)	e.g., measured spectral reflectances or radiance, radar reflectivity, etc.
X	state quantities (n elements)	unknowns to be retrieved
$F(\mathbf{X})$	forward model (m)	maps X to Y, e.g., radiative transfer model using ancillary datasets and/or assumed model parameters
X$_a$	state *a priori* (n)	e.g., climatology
X$_{op}$	optimal state given the observations (n)	the solution of the 'retrieval'
B	model parameters (k)	
Matrices		
S$_y$	covariance of **Y** (m x m)	Uncertainty in the observation
S$_a$	covariance of **X$_a$** (n x n)	Uncertainty in the *a priori*
S$_b$	covariance of **B** (k x k)	Uncertainty in the model parameters
S$_e$	covariance of **Y** + covariance of **B** mapped into observation space via **K$_b$** (m x m)	Uncertainty in the total measurement (observational uncertainty plus simulated observation by the forward model). This assumes no correlation between the Y and B error sources
S$_{op}$	covariance of **X$_{op}$** (n x n)	Uncertainty in the retrieved solution
K$_b$	Jacobian of F w.r.t. **B** (m x k)	Elements $\partial F_i / \partial B_j$ given **X$_{op}$** and **B**
K$_{op}$	Jacobian of F w.r.t. **X$_{op}$** (m x n)	Elements $\partial F_i / \partial Xop_j$ given **X$_{op}$** and **B**

3
4

1

Forward Model
F

Geophysical Variable
(What we want to know)

Radiance or Backscatter
(What we observe)

X

$X + \delta X$

$\hat{X}$

Y

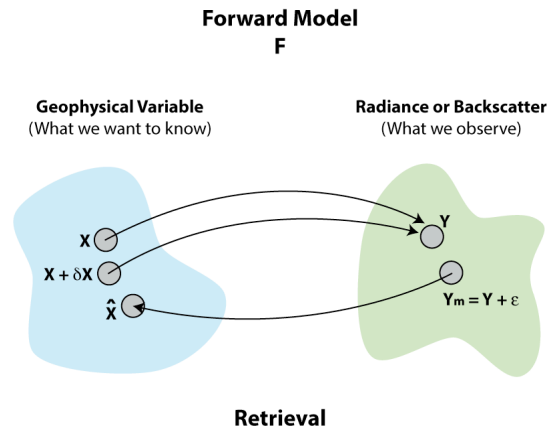$Y_m = Y + \varepsilon$

Retrieval

Figure 1: Illustration of the forward model / retrieval problem. Note that the geophysical variable space and observational space are typically multi-dimensional (i.e., we desire to retrieve n geophysical variables from m observations, where n typically does not equal m).
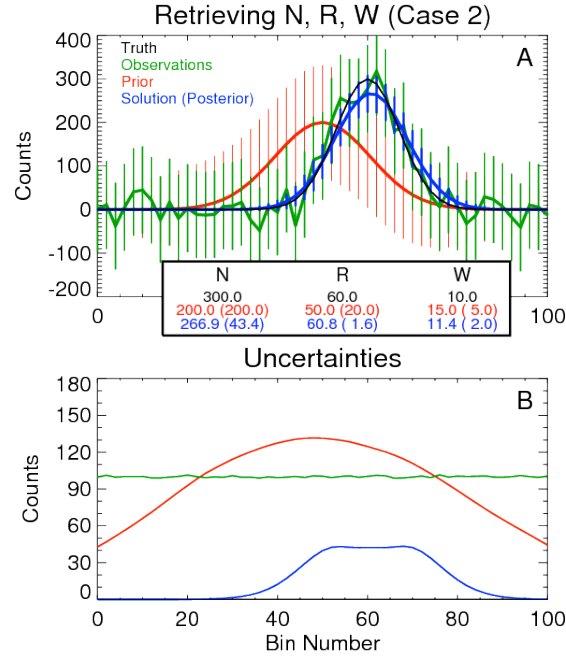
2

3

1

2

Figure 2: The "perfect model" case in where there are no unknown parameters; the entire model is described by the state vector **X**=[N, R, W]. Panel A shows the true size distribution (black), the noisy observations (green), the prior distribution (red), and the retrieved solution (blue), where the latter two cases are translated from the state vector space to observational space by *F*. The error bars are the 1-σ uncertainties that are derived from a Monte Carlo sampling of the S_e, S_a, and S_{op} matrices, respectively. The values in the inset box indicate the true, prior, and retrieved values of the three state vector parameters, with the 1-σ uncertainties in parentheses. Panel B shows the 1-σ uncertainties, again in the observation space, derived from the Monte Carlo sampling of S_e (green), S_a (red), and S_{op} (blue). Since there are no unknown parameters in this perfect model, $S_e = S_y$ and S_b=0, which is why the green line in panel B is flat.
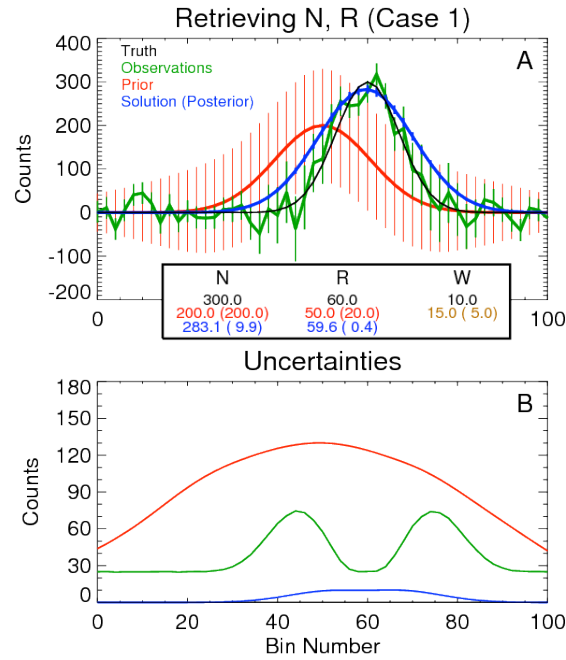
3

Figure 3: Similar to Figure 2, but using an imperfect model where W is a model parameter with some uncertainty and the state vector being retrieved is **X**=[N,R]. The model parameter and its 1-σ uncertainty are in brown in the inset box. The uncertainty in the observations (**S$_y$**) is also 4x smaller than that used in Fig 2 (i.e., the same uncertainty as in online supplemental Fig S1). Note that the uncertainty in the observations (derived from the **S$_e$** matrix) now has contributions from both **S$_y$** and **S$_b$**.

2

3

4
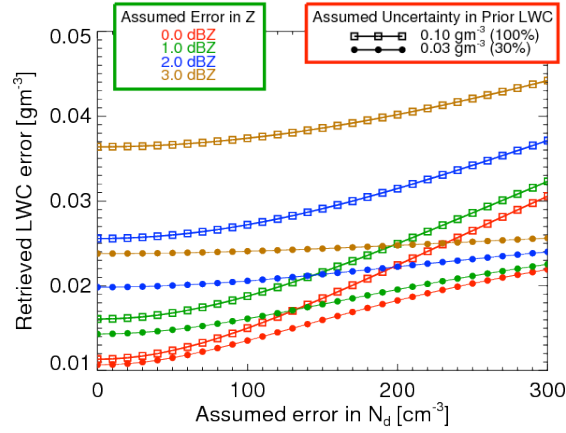
Fig 4: The uncertainty in the retrieved *LWC* as the uncertainty in the model parameter N_d (x-axis), radar reflectivity observation uncertainty (green box), and prior uncertainty (red box) are changed. The retrieved *LWC* was 0.1 g m^{-3} using N_d = 500 cm^{-3} and **Y**=-37 dBZ.

1

2

1

2

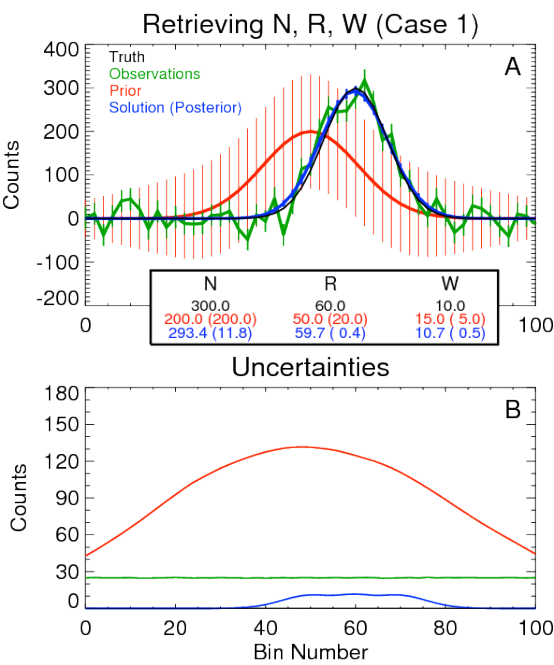**Retrieving N, R, W (Case 1)**

Figure S1: The same as Figure 2, except that the uncertainty in the observations (i.e., $\mathbf{S_y}$) was assumed to be 4x smaller.

3

4

5
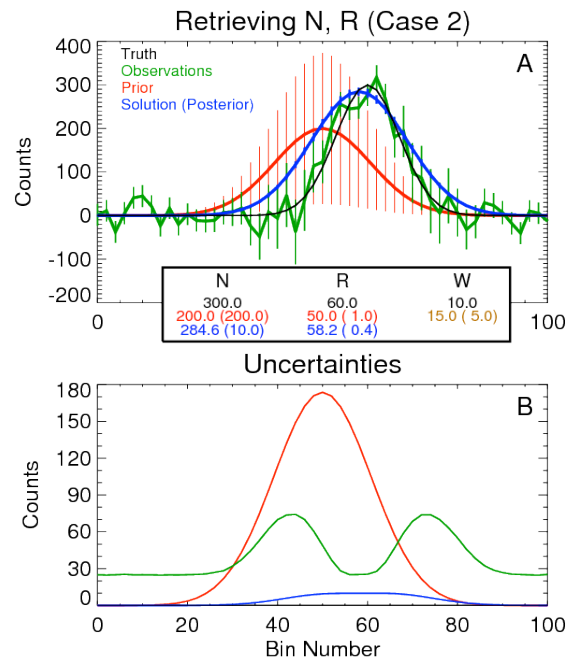
Figure S2: The same as Figure 3, except that the uncertainty in R in the prior is 20x smaller.
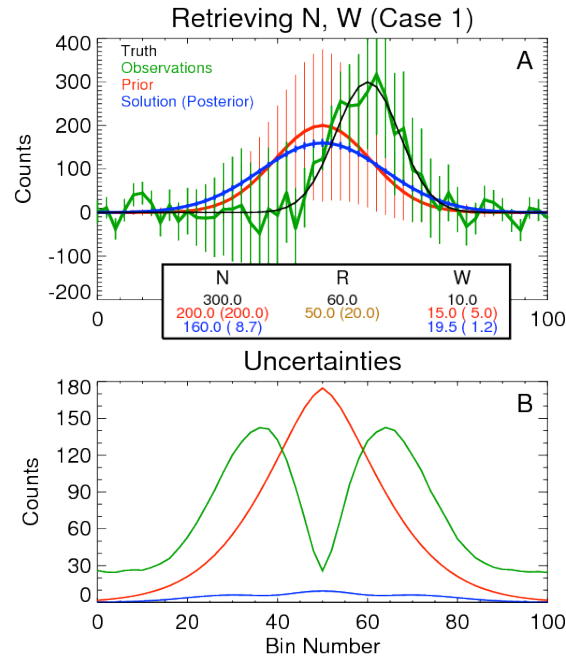
1

2

1

2

3

Figure S3: Similar to Figure S2, but using an imperfect model where R is the model parameter with some uncertainty and the state vector being retrieved is **X**=[N,W]. Note that the uncertainty in the observations (derived from the $\mathbf{S_e}$ matrix) again has contributions from both $\mathbf{S_y}$ and $\mathbf{S_b}$.
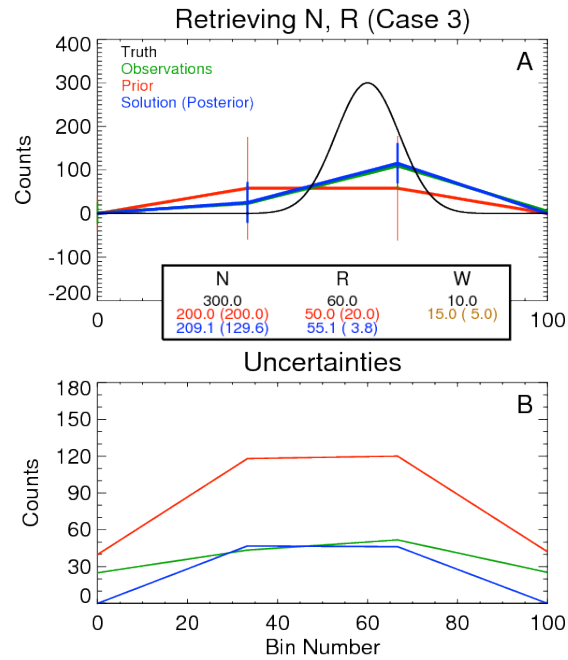
4

Fig S4: The same as Fig 3, except that the number of observations was decreased from 50 to 4 (i.e., from an observation every 2nd bin to every 33rd bin). In this example, the observations are almost exactly matched by the solution (which was translated into the observational space).

1

2